Computers, Materials & Continua
DOI:10.32604/cmc.2021.015647
Article

Tech Science Press

Deep Learning-Based Hybrid Intelligent Intrusion Detection System

Muhammad Ashfaq Khan[1,2] and Yangwoo Kim[1,*]

[1]Department of Information, and Communication Engineering, Dongguk University, Seoul, 100-715, Korea
[2]Department of Electronics Engineering, IoT and Big-Data Research Center, Incheon National University, Incheon, Korea
[*]Corresponding Author: Yangwoo Kim. Email: ywkim@dongguk.edu

Received: 01 December 2020; Accepted: 24 January 2021

Abstract: Machine learning (ML) algorithms are often used to design effective intrusion detection (ID) systems for appropriate mitigation and effective detection of malicious cyber threats at the host and network levels. However, cybersecurity attacks are still increasing. An ID system can play a vital role in detecting such threats. Existing ID systems are unable to detect malicious threats, primarily because they adopt approaches that are based on traditional ML techniques, which are less concerned with the accurate classification and feature selection. Thus, developing an accurate and intelligent ID system is a priority. The main objective of this study was to develop a hybrid intelligent intrusion detection system (HIIDS) to learn crucial features representation efficiently and automatically from massive unlabeled raw network traffic data. Many ID datasets are publicly available to the cybersecurity research community. As such, we used a spark MLlib (machine learning library)-based robust classifier, such as logistic regression (LR), extreme gradient boosting (XGB) was used for anomaly detection, and a state-of-the-art DL, such as a long short-term memory autoencoder (LSTMAE) for misuse attack was used to develop an efficient and HIIDS to detect and classify unpredictable attacks. Our approach utilized LSTM to detect temporal features and an AE to more efficiently detect global features. Therefore, to evaluate the efficacy of our proposed approach, experiments were conducted on a publicly existing dataset, the contemporary real-life ISCX-UNB dataset. The simulation results demonstrate that our proposed spark MLlib and LSTMAE-based HIIDS significantly outperformed existing ID approaches, achieving a high accuracy rate of up to 97.52% for the ISCX-UNB dataset respectively 10-fold cross-validation test. It is quite promising to use our proposed HIIDS in real-world circumstances on a large-scale.

Keywords: Machine learning; intrusion detection system; deep learning; spark MLlib; LSTM; big data

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

The intrusion detection (ID) system is a renowned solution for detecting malicious activities in a network. The types of malicious network attacks have grown exponentially, and the ID system has become an essential component of defense in addition to network security infrastructure. In 1931, John Anderson published the first significant paper on ID, Computer Security surveillance, and threat monitoring [1]. An ID system usually monitors all internal and external packets of a network to detect whether a packet has a sign of intrusion. A well-made ID system can determine the properties of numerous malicious activities and automatically respond to them by sending cautions.

In general, there are three common ID system classes; these classes are based on detection approaches. The first class is the signature-based system (SBS), which includes the misuse detection technique. The second is the anomaly-based system (ABS), also known simply as "anomaly." The third one is the stateful protocol analysis detection [2]. SBS relies upon a pattern matching technique, taking a database of known attack signatures and comparing these to signatures present in the observed data. An alarm goes off when a match is identified. SBS detects attacks based on existing knowledge; as such, the misuse detection technique is also recognized as a knowledge-based technique. The misuse detection technique features a minimum false alarm rate and maximum accuracy; it cannot, however, identify strange attacks. Similarly, the behavior-based ID system, also known as ABS, can detect intrusion by matching normal behavior to an abnormal one. The stateful protocol ID method compares the known malicious activities and identifies the eccentricity of protocol activity, taking advantage of both anomaly and signature-based ID techniques. ID systems can be further categorized into three types according to their architectures: Network-based detection system (NIDS), Host-based detection system (HIDS), and the hybrid approach [3]. For a HIDS application, the software is fixed, and the host computer plays an important role in evaluating and monitoring system behavior and event log files play active roles in ID [4]. Unlike a HIDS, which analyzes each host separately, a NIDS analyzes the packets that flow above the network. This gives the NIDS an edge over the HIDS because it can test the whole network with a unique system structure. However, while the NIDS is superior in terms of installation cost and time of application software, it is vulnerable to distribution into a system over the network and affects the complete network. The hybrid IDS combine both the HIDS and NIDS with better-quality security mechanisms. The hybrid system joins the spatial sensors to identify vulnerabilities, which can occur at a particular point or over the whole network. There are two main ID system types, which are defined according to the system's deployment structure: distributed structure and non-distributed structure. A distributed structure involves several ID subsystems that communicate with each other over an extensive network. In contrast, a non-distributed system can be mounted only at a single, unique location, for example, an open-source snort.

Most approaches currently used in ID systems are unable to deal with the complex and dynamic nature of malicious threats on computer networks. Therefore, effective adaptive methods, such as several ML techniques, can achieve a higher intrusion detection rate (DR), a low false alarm rate (FAR), reasonable communication, and computation costs. There are various customary approaches to intrusion identification, including access control, cryptography, and firewalls. These traditional ID techniques have few limitations in fully protecting a system; most notably, when systems are facing a high volume of malicious attacks, DOS and systems can obtain high values of FP and FN attack DR. Recently, numerous researchers have used ML techniques for ID to improve ID rates. Several studies have been done to enhance and apply this method to the ID system. In the present study, we also reviewed several articles that use state-of-the-art ML methods for ID. The ML models were found to have many issues that slow down the training process; these issues included the size of the dataset and the optimal parameters for the most suitable model. These kinds of problems prompted the researchers to look for the most effective methodology. The use of open source clustering computing, such as Spark (a big data analytics system), is one potential solution to such problems. We propose a unique tactic to improve ID system performance. However simple ML approaches are limited, while intrusion methods are expanding and growing increasingly complex. Advanced learning approaches are essential, especially in the analysis of intrusion and feature extraction. Hinton et al. [5] stated that DL has attained great achievements in domains such as image processing, NLP, and weather prediction.

With the growth of cyber-security capabilities, cyber-attacks have risen to the challenge of breaching new security defenses. Considering the possibility of simultaneous attacks, it is vital to select an appropriate action by proactively predicting and evaluating the effects of a specific security event. Cyber-attacks, mainly in the sphere of large-scale military networks, can have a lethal influence on security; therefore, numerous tests and extensive research are required preparations. Today, cyber-space is known as the fifth battlespace, following land, air, space, and sea; cyber warfare can affect the military strategies and activities that are associated with national security. Although the military is working to recognize and minimize cyber-attacks, cyber-attacks are consistently on the rise [6,7]. It is significant to observe the malicious threat that arises in distinct ways to respond effectively to it. The significance of cyber-attacks on the infrastructure that should be protected and security policies established. It is not only to analyze cyber threats but also to increase the possibility of more proactive reactions. Numerous studies have been carried out on cyber-attacks modelings, such as the attack graph, attack tree, and cyber kill chain modeling approach [8,9]. Notice that previous works on cyber-attack modeling were limited in a large-scale network environment, due to problems such as scalability. Nowadays, cyber threats do not stop with a single attack but come in complex forms that involve numerous kinds of cyber-attacks. Besides, novel attacks are constantly emerging. To overcome these challenges, a novel approach to modeling that is flexible enough to adapt to new attacks easily and systematically is required [10].

As mentioned above, misuse and anomaly ID methods have their limitations. Our proposed HIID approach combines the two approaches to overcome their respective shortcomings while maintaining their advantages, which involve improved performance compared to conventional techniques. To increase IDS learning ability and performance, we propose a better-quality ID system that consists of Spark MLlib and state-of-the-art DL approaches, such as LSTMAE. The key contributions of our research may be summarized as follows:

- The development of HIIDS, which relies on Spark MLlib and state-of-the-art DL techniques, such as LSTMAE, which merges both shallow and deep networks to overwhelm their analytical overheads and exploit their benefits. This HIIDS investigates how to solve the class imbalance problem that usually occurs in ISCX ID datasets.
- Further investigation of the packet capture file directly on Spark; prior studies did not evaluate the raw packet dataset.
- Comparison of the HIIDS with other conventional ML methods. The simulation results demonstrate that the HIIDS approach is highly appropriate for malicious traffic detection. It has higher attack detection accuracy and was found to correctly detect network misuses in 97.52% of the cases through 10-fold cross-validation.

The rest of this paper is organized as follows. The background of ID and related work are briefly reviewed in Section 2. A brief overview of our proposed HIIDS and a detailed description of the dataset that was used for classification are provided in Section 3. A simulation of our proposed framework with performance metrics is discussed in Section 4. The paper is concluded with a possible direction for future work described in Section 5.

2 Related Work

Over the last two decades, the application of machine learning (ML) and deep learning (DL) to intrusion detection (ID) systems has been suggested by several researchers. Therefore, various models have been developed for network intrusion detection (NID) using conventional ML techniques. Examples include K nearest neighbors (KNN) as suggested by Khammassi et al. [11] Logistic Regression (LR) as suggested by Moustafa et al. [12] Support vector machine (SVM) as suggested by Khan et al. [13] Random Forest (RF) as suggested by Farnaaz et al. [14] Decision Tree (DT) as suggested by Sindhu et al. [15] Naïve Bayes (NB) as suggested by Buczak et al. [16] and Artificial Neural Networks (ANN) as suggested by Vincent et al. [17]. However, these prior techniques demonstrate inadequate classification performance with the maximum false alarm rate (FAR) and low attack detection rate (DR) in an ID system. Kim et al. [18] developed a hybrid system that incorporates misuse and anomaly using supervised ML classifiers SVM and DT, respectively, and assessed their hybrid approach using NSL_KDD older data. The authors claimed that the improved attack detection accuracy was owing to the hybrid ID system. Paulauskas et al. [19] developed a novel approach for ID using various weak learners; this is known as the ensemble approach. The weak learners have low malicious detection accuracy. There were some weak learners, such as J48, C5.0, Naïve Bayes, and rule-Based classifiers that were used by the authors. Zaman et al. [20] used a better-quality ID algorithm recognized as enhanced support vector decision function (ESVDF) and evaluated their proposed IDS using the DARPA dataset; the proposed IDS was found to be superior to other conventional ID approaches.

Table 1: Summary of the related works using different approaches

Reference	Approach	Accuracy (%)	Dataset
Tang et al. [21]	DBN + LR	97.0	KDD99
Qatf et al. [22]	SAE + SVM	93.96	KDD99
Qatf et al. [22]	Deep VAE	84.96	NSL_KDD
Farahnakian et al. [23]	AE	94.71	KDD99
Naseer et al. [24]	DNN	89.0	NSL_KDD
Bandyopadhyay et al. [25]	DCNN	84.58	NSL_KDD
Albahar et al. [26]	Ensemble	93.3	UNSW-NB15
Monshizadeh et al. [27]	MCA + EMD	87.29	ISCX-2012
Thi-Thu et al. [28]	FS + DT	95.33	ISCX-2012
Mighan et al. [29]	SAE + SVM	90.3	ISCX-2012
Wang et al. [30]	HAST − IDS	96.6	ISCX-2012

Although the above ID approaches have demonstrated decent accuracy up to a certain level, certain improvements, such as decreasing the number of FAR and increasing the ID accuracy, are necessary. In this regard, DL is a powerful technique. DL is a branch of ML that has become progressively dominant in various fields, such as speech recognition and natural language processing (NLP). DL's popularity is due to its two fundamental characteristics: (a) hierarchical features representations and (b) handling of long-term dependencies of sequential patterns. Today, state-of-the-art DL approaches that are used for NID includes auto-encoders (AE), deep belief networks (DBNs), deep neural networks (DNNs), and restricted Boltzmann machines (RBMs) as well as variants of these approaches. An overview of the state-of-the-art approaches is presented in Tab. 1.

DL has shown that its attack detection accuracy in the ID domain effectively exceeds conventional approaches [31]. Erfani et al. [32] developed a novel tactic that joined one-class linear SVM with the DBN for ID, evaluating it with various benchmark ID data. Fiore et al. [33] proposed a new technique, discriminative RBM, to learn compressed attributes from data attributes; these compressed attributes are then used for binary classification purposes into softmax classifier for benign and malicious network behaviors. Wang et al. [34] presented a DL-based IDS based on AE for detecting network traffic from the raw dataset and achieved a very high ID performance. Javaid et al. [35] used the DNN technique for anomaly detection. Their evaluation based on DNN DL found that the DNN technique is a novel and effective approach for ID in a software-defined network (SDNs). Yin et al. [36] introduced a neural network (NN) DL-based NIDS. This DL-based ID was tested using the NSL_KDD dataset; it was found that the DL-based IDS outperformed conventional ML-based classification techniques. Khan et al. [37] presented the hybrid DL approach for ID and applied it to real-time ID data. Their simulation outcomes showed that the hybrid DL-based IDS was superior in terms of attack classification accuracy and performance. Alrawashdeh et al. [38] proposed a DL-based IDS using the DBN of RBM with four and one hidden layers for attribute reduction purposes; the weights of the DBN were restructured during fine-tuning, and attack classification was accomplished using an LR classifier. The developed methodology was evaluated on the benchmark KDD99 data and attained an attack classification accuracy of up to 97.9% with a FAR of 0.5%. However, the attack classification accuracy as evaluated using this ancient data is not sufficient to show that this a robust approach for NID. Shone et al. [39] proposed a non-symmetric deep AE-based ID and evaluated the proposed framework with the benchmark KDD99 dataset, achieving an attack classification accuracy of 97.87% and a FAR of 2.15%. In [40], the authors aimed for a Deep Neural Network (DNN) of 100 hidden units. To improve performance, they utilized a GPU and the KDD99 dataset. The authors proposed that the models of both recurrent neural network (RNN) and long short-term memory (LSTM) are better for enhancing the attack detection accuracy. These ID systems based on DL techniques were found to be superior to traditional approaches; the authors also presented various ideas by joining DL and ML techniques, with the primary goal of developing an efficient and robust ID system. Wang et al. [41] developed a novel approach for ID by combining fuzzy clustering and ANN; they tested a novel hybrid approach on the KDD99 dataset and demonstrated that their hybrid FC-ANN approach outperforms traditional ML approaches in terms of ID. Mukkamala et al. [42] used a hybrid approach by combining the SVM and ANN; they evaluated this approach on the benchmark KDD99 dataset. Here, SVM and ANN were used for classification tasks and data patterns, respectively. Various researchers have used the ISCX-2012 ID dataset to conduct suitable system validation. However, there is still much room for enhancements, such as improving attack detection accuracy and reducing FAR [43–48]. ID research has been carried out by various scholars for developing both the ABS and SBS using separate classification methods. These methods fail to afford the efficient possibility of attack detection, so a hybrid ID system is an important research challenge. ML-based techniques have been mostly used by scientists and engineers to develop an ABS, which can make a model by comparing normal with abnormal behavior and then attempting to classify whether upcoming new packets are "attack" or "normal." DL is enormously valuable for the ID system because it automatically extracts features of the specific problem without requiring robust preceding knowledge. The main downside of using the DL model for the ID domain is the extent of the training; obtaining the right model is time-consuming.

The research community has drawn substantial attention to the issue of class imbalance [49]. The problem of class imbalance is created by insufficient data distribution; one class contains most samples, while others contain comparatively few. The classification problem becomes more complicated as data dimensionality increases due to unbounded data values and unbalanced classes. Bedi et al. [50] utilized numerous ML approaches to deal with the class imbalance issue. Thabtah et al. [51] also evaluated various approaches to the class imbalance problem. Most data samples are targeted by most of the algorithms while missing the minority data samples. As a result, minority samples appear irregularly but constantly. The main algorithms for solving the unbalanced data problem are data preprocessing and feature selection techniques, and every approach has both benefits and shortcomings. The ID dataset has a high-dimensional imbalance problem including missing features of interest, missing feature values, or the sole existence of cumulative data. The data appear to be noisy, containing errors and outliers, and unpredictable, comprising discrepancies in codes or names. We used over-sampling to resolve the problem of the imbalance; this involved enlarging the number of instances in the minority class by arbitrarily replicating them to increase the presence of the minority class in the sample. Although this procedure has some risk of overfitting, no information was lost, and the over-sampling approach was found to outperform the under-sampling alternative.

With the accelerated growth of big data, DL approaches have flourished and have been widely utilized in numerous domains. In contrast to previous studies, we took a hybrid approach—the Anomaly-Misuse ID method—to two-stage classification to overwhelm the condition face by separate classification methods. We used Spark MLlib and the LSTMAE DL approach for ID, on the well-known real-time contemporary dataset ISCX-2012.

3 Proposed Approach

Fig. 1 presents the anticipated ID framework. It comprises two learning stages. For this HIID, we planned to construct a two-stage ID system, in such a way that Spark MLlib as an anomaly in Stage-1 and LSTMAE as misuse in Stage-2.

These two stages of ID framework are efficient in terms of computational complexity while using full features datasets and offer a higher accuracy with a low probability of FAR.

Stage-1 Anomaly detection using Spark MLlib classifiers.

Stage-2 Misuse detection using state-of-the-art deep learning approaches such as LSTMAE.

3.1 Overview of the Proposed IDS

The hybrid framework concentrates on resolving real-time ID problems, using enormous data analysis models (Apache Spark and Apache Hadoop) and AI (ML and DL). Controlling such an issue is a complex task due to space and time restrictions. Big data is enormous and consistently increasing in volume but requires prohibitive amounts of power, specialized resources, and a computational device that can effectively handle the data. The hybrid ID framework overcomes these problems by using the MLlib with LSTMAE. The main structure of the HIIDS existing here, forms the source of the experiment, to use Spark MLlib and deep learning.

Numerous ML techniques were used due to the huge volume of data. We selected the competent Spark to implement a logistic regression (LR) and extreme gradient boosting (XGB) classifiers. Initially, preprocessing data was delivered through these machine learning classifiers to produce regression models that present the opportunity of all data. Generally, this is the binary learning phase. In this hybrid ID approach, the NIDS using both anomaly and misuse techniques. The proposed hybrid ID architecture contains a data preprocessing module, a Spark MLlib classification component integrating the anomaly detection module (Stage-1) of the proposed hybrid IDS with misuse detection, and DL classification (Stage-2), followed by the alarm module.

Stage-1 utilized Spark MLlib to perceive anomalies that may be intrusions, and Stage-2 utilized the LSTMAE DL model, which further classifies attacks in the event they occur. The details of the proposed Spark MLlib and LSTMAE model are shown in Fig. 1.

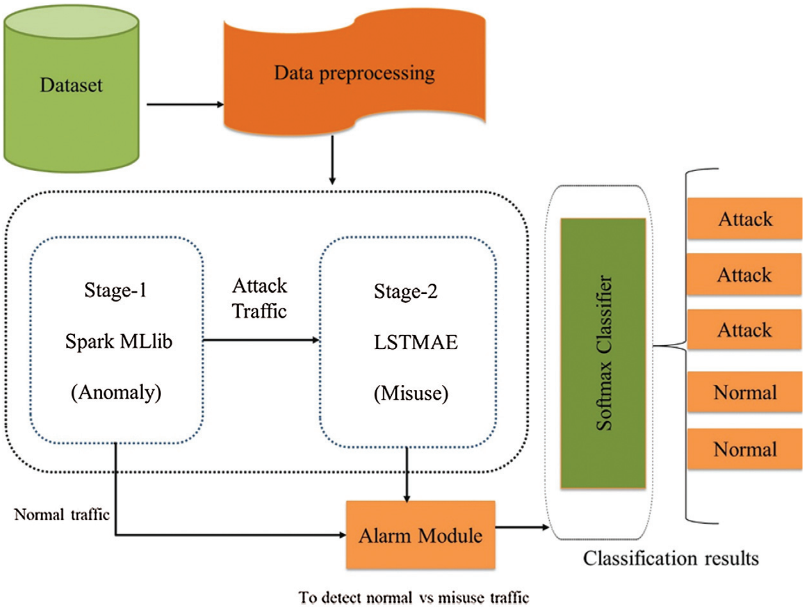

Figure 1: The micro overview of the proposed ID framework

The architecture of the hybrid IDS is as shown in Fig. 1; initially network traffic was arranged and preprocessed. During preprocessing, all necessary conversions were made for both Stage-1 Spark MLlib and Stage-2 LSTMAE-based modules of HIIDS; both stages had their supported data formats. For our hybrid ID experiment, we used 1,512,000 network traffic packets attained from ISCX-2012 datasets to demonstrate the effectiveness of the proposed HIIDS.

3.2 Datasets

Choosing a suitable ID dataset plays a significant role in testing the ID system; therefore, the simulation of the proposed HIID approach was carefully deliberated.

3.2.1 Explanation of the ID Dataset

There are various standard datasets, and some of them comprise inflexible, outdated, and irreproducible attacks. To overcome these shortcomings and create further up-to-date traffic patterns the ISCX-2012 data was created by the Canadian institution of cybersecurity [52]. It contains various types of ID data to assess anomaly ID approaches. The ISCX-2012 data shows real network activities and includes numerous attack scenarios. Additionally, it is shared as a complete network capture with completely internal network traffic to assess payloads for network packet analysis. The ISCX 2012 ID data comprises both malicious and normal traffic actions for seven consecutive days. The data was created by profiles containing abstract representations of network traces' particular behaviors and activities.

Communication between the destination and source host over HTTP can be represented by the sending and receiving of packets, endpoint attributes, and other similar features. This illustration produces a unique profile. It produces realistic network traffic for POP3, SSH, FTP, HTTP, SMTP, and IMAP protocols.

The ISCX-2012 contains dual distinct profiles to make network traffic activities and states. The multi-stage or abnormal states of abnormal attacks are identified by an α profile, while feature characterization and mathematical dissemination of the method are done with the β profile. For example, the β profile can comprise network traffic packet size distributions in the explicit patterns and time distribution request of the protocol, whereas the α profile is created based on the sophisticated preceding attacks of a distinct day. The α profile consists of four kinds of attack scenarios.

(1) Internal infiltration of the network traffic

A vulnerable application program, such as Adobe Acrobat Reader, generally takes the advantage of the internal infiltration of a network. A backdoor can be performed on the victim's machine after successful penetration and will execute several malicious attacks on the victim's network. To detect these kinds of malicious threats, mostly applied Nmap and port scan.

(2) HTTP DOS attacks

The attacker causes a network resource to be unavailable for a particular time. This is typically done by overwhelming a network resource with superfluous requests to overwork the network and impede the fulfillment of some or all legitimate requests. To collect these kinds of DoS attacks, mostly utilized the Slow HTTP test, Hulk, Slow loris, and Goldeneye.

(3) DDOS using IRC botnet

These types of attacks generally occur when various networks flood the bandwidth or various resources of a particular victim. Therefore, a DDoS attack is often the result of various infected networks (for example, a botnet), which flood the target network by creating massive network traffic. These types of attacks have utilized LOIC for UDP, TCP, and HTTP.

(4) Brute force SSH

This is the most common type of attack that can be used not only to crack passwords but also to determine secret content and pages of several web applications. These types of attacks have been launched via FTP and SSH Patator tools.

The full ISCX-2012 dataset is shown in Tab. 2. It can be observed that each attack state was realistic for a single day, while two daysn consisted of normal traffic. The variety of normal behavior and the complication of malicious attack states in the network have been previously described [52].

3.2.2 Data Preparation and Feature Engineering

The dump network traffic was initially preprocessed and prepared, as shown in Fig. 1. The ISCX-2012 dataset was evaluated, and after preprocessing, it was composed of seven consecutive days with the systematic and practical circumstances reflecting network attacks. The data were labeled for malicious and benign streams for a full of 68,792, and 2,381,532 records in the own class. The abnormal attacks were detected in the initial traffic data and were divided into two classes: benign/normal and abnormal/malicious.

Table 2: Daily traffic ISCX-IDS 2012 dataset summary

Days	Date	Explanation	Size (GB)
Sunday	13/6/2010	Infiltrating the traffic from internal and regular activities	3.95
Monday	14/6/2010	HTTP DOS and regular activities	6.85
Tuesday	15/6/2010	DDOS with a Botnet IRC	23.04
Wednesday	16/6/2010	Normal, hence there are no abnormal activities	17.6
Thursday	17/6/2010	Brute force (SSH) and regular activities	12.3
Friday	11/6/2010	Regular, hence there are no abnormal activities	16.1
Saturday	12/6/2010	Infiltrating the Network traffic from internal and usual activities	4.22

Furthermore, several multi-stage malicious intrusion scenarios were executed to generate various attack traces (e.g., HTTP, DoS, brute force SSH, infiltration from the interior, DDoS via an IRC botnet). The detailed descriptions of training and testing data distributions are presented in Tabs. 3 and 4.

Table 3: Testing and training data distribution of ISCX-2012

Network flows	#Features	Testing		Training	
		Benign	Malicious	Benign	Malicious
ISCX-UNB Saturday	8	45,889	1,353	85,222	1,353
ISCX-UNB Monday	8	58,664	1,320	108,945	2,451
ISCX-UNB Tuesday	8	187,012	13,083	347,308	24,295
ISCX-UNB Wednesday	8	182,793	0	339,470	0
ISCX-UNB Thursday	8	137,338	1,822	255,054	3,381

The core idea of this research was to evaluate the reliability of the hybrid system against anomalies and the unknown, via the misuse approach. Tab. 4 presents the testing and training network traffic data for misuse attack detection using a state-of-the-art DL approach, such as an LSTMAE.

Table 4: ISCX-2012 data distribution for stage-2

No. of input	No. of features	Attack type
Train set	8	DDoS, Brute force SSH, HTTP DoS, and infiltration traffic from the inside
Test set	8	DDoS, Brute force SSH, HTTP DoS, and infiltration traffic from the inside

3.3 Implementation Details

The dominance of the HIIDS is evaluated through experiments applying the ISCX-2012 ID datasets via normal and attack classifications: false positive, false negative, true positive, attack detection precision, and error rate. To show the efficacy of our suggested ID system, we executed the first stage in Scala by Spark MLlib for anomaly detection; the second stage was executed for misuse detection; the DL approach was executed in Java with Deeplearning4j. The simulation was done on a 64-bit cluster computer with 32 cores, 32 GB RAM, and Ubuntu version 14.04 OS. The software stack contained Java (JDK) 1.8, Spark v2.3.0, Deeplearning4j 1.0.0. alpha, and Scala 2.11.8. The deep learning was trained on an RTX 2080 Ti GPU with cuDNN, and CUDA facilitated the pipeline speed.

To measure the HIIDS performance, we first split the dataset into train and test datasets. To form an efficient HIID framework, we utilized the training data and analyzed our hybrid approach with testing data. The block diagram of our anticipated HIID is presented in Fig. 1. The ISCX-2012 with complete, original features are utilized to demonstrate the dominance of our proposed hybrid approach. The network traffic mixed with malicious and normal pass through spark MLlib Stage-1 which categorized data into malicious and normal classes. Stage-2 LSTMAE was modeled with malicious traffic; malicious traffic was further categorized into 4 analogous attacks. The hybrid approach overcomes the computational complexity while applying comprehensive features to the ISCX-2012 dataset with higher ID accuracy and low FAR. 80% of the data with 10-fold cross-validation was utilized for training purposes, and the model was evaluated with a 20% held-out dataset.

3.3.1 Stage-1: The Anomaly-Based Detection Module

Apache Spark is a competent big data processing engine for detecting cybersecurity attacks. Spark MLlib is the most efficient big data analytics library currently available, executing over 55 ML algorithms [53,54]. Spark MLlib is most suitable for ML tasks and is 10 times faster than Hadoop-based big data processing tools for iterative tasks. MLlib of spark evolution was initiated in 2012 as a portion of an ML-based project, and in 2013 it became an effective open-source library for ML tasks. Spark MLlib contains several ML algorithms for instance classification, clustering algorithms, and regression and dimensionality reductions that are crucial to the development of classic ML real-time applications; its mechanisms have been established by several scholars to progress high dimensional data analytics worldwide.

MLlib-based anomaly attack detection at Stage-1 was first modeled based on an established training set, which contains both normal and malicious traffic. The test data that contain unknown, regular, and malicious traffic are used to validate the anomaly module of IDS. The attack observed on original traffic data were divided into two classes: Abnormal (malicious) and normal. Abnormal network traffic behavior is known as anomaly traffic. Detection of this kind of abnormal network traffic was passed through the Stage-2 that LSTMAE, where the misuse attack detection technique did further attack detection and classification.

3.3.2 Stage-2: A Misuse Detection Module

LSTMAE was used in this stage to define the misuse of network traffic and goals of further classifying the anomalous traffic according to specific policies. An overview of the misuse detection module using an LSTMAE is given in Fig. 2. LSTM is an upgraded version of the RNN, which was introduced in [55,56] to efficiently address vanishing and exploding gradient issues. All hidden layers of RNN are substituted with memory blocks that comprise a memory cell intended to reserve information, with three important gates that play dynamic roles in LSTM (Input, Output, forget gate) [57]. The most powerful feature of LSTM lies in its capability to capture long dependencies and learn competently from variable amount sequences.

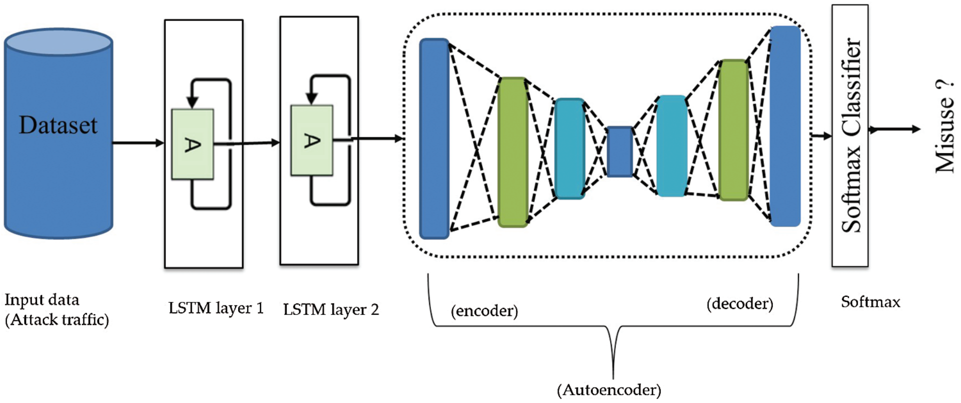

Figure 2: The micro-overview of LSTMAE

Research has shown that LSTM demonstrates high confidence and effectiveness for resolving issues of video classification [58], sentiment analysis [59], emotion recognition [60], and abnormal activities [61].

In this module, LSTMAE is used as a misused attack detection technique. LSTMAE misuse attack detection techniques aim to further categorize the abnormal data from Stage-1 among equivalent classification: DOS, Scan, HTTP, and R2L. While misuse ID uses the LSTMAE, the technique was initially trained in the abnormal traffic to create a model that provides the baseline profile for abnormal traffic. A test set is an input to the training model that tests whether the training model performance is malicious (abnormal) or normal. An alarm goes off when a match is found. More internal information can be effectively obtained with LSTMAE, compared with other hand-crafted techniques.

4 Experimental Evaluations

A detailed description of the experimental results will be discussed in this section. Since the dominance of the proposed HIID is sensibly analyzed, this can only be realized throughout experiments applying the ISCX 2012 ID datasets via normal and attack classification, false positive, false negative, true positive, attack detection accuracy, and error rate.

4.1 Performance Metrics

The elements of the confusion matrix that assist in representing the expected and predicted classification are given in Tab. 5. The outcome of classifying is predicated among two-class issues such that correctly and incorrectly. Four essential states must be computed in the confusion matrix.

- True Positive (TP). It presents that model is accurate as normal and predicts positive and it is represented by x.

- False-negative (FN). It represented the wrong prediction and denoted by y. It identifies instances that are malicious in certainty, as normal and the model inaccurately predicts negative.
- False-positive (FP). It presents a model that mistakenly predicts positive and, the number of detected attacks is normal. It is represented by z.
- True negative (TN). It is represented by t and specifies instances that are correctly observed as an attack predicts negative.

Table 5: Confusion matrix for proposed IDS

	Predicated		
Actual	Normal	TP	FN
	Anomaly	FP	TN

From the above-mentioned conditions of the confusion matrix, we can compute the performance of the system as follows. The two most essential and general parameters for the evolution of the ID system are TPR or DR and FAR. The percentage of intrusion instances recognized by the ID model is known as DR, while the amount of misclassified normal instances is known as FAR.

$$\mathrm{TPR} = \mathrm{DR} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) = \mathrm{x}/(\mathrm{x} + \mathrm{y}) \tag{1}$$

$$\mathrm{FAR} = \mathrm{FP}/(\mathrm{TN} + \mathrm{FP}) = \mathrm{z}/(\mathrm{t} + \mathrm{z}) \tag{2}$$

We claim that the HIIDS is superior to conventional IDS, as it increases DR and decreases FAR.

4.2 Evaluation of the Hybrid IDS

Tab. 6 presents the overall performance of several classifiers. The results of the random search are described in this section. As presented in the table, the classical LR model gave an F1-score accuracy of ∼83%, whereas tree-based ML classifiers managed to considerably increase the accuracy to 88%.

However, the most significant improvement that we observed was with state-of-the-art DL approaches such as LSTMAE, which correctly identified misuse for up to 97.0% of cases. This improvement was due to the temporal feature's extraction with LSTM and the extraction of more important internal information by the AE.

Table 6: Classifier performance at several stages

Classifier	Stage	Precision	Recall	F1-score	FAR	DR
LR	1	0.830	0.823	0.8264	10.50	0.82
XGB	1	0.8775	0.8745	0.8759	8.13	0.87
LSTMAE	2	0.9653	0.9752	0.9702	1.2	0.9752

4.3 *Overall Analysis*

Tab. 7 summarizes the results of the current approach for the ISCX-2012 data. These datasets were produced later than the KDD and DARPA data, so only a few corresponding tentative results exist. Therefore, using the existing simulation results, the best outcomes for each stage are defined by FAR and accuracy. It is evident that the proposed ID system performs well, both in terms of accuracy and FAR related to state-of-the-art techniques. This is owing to the Spark MLlib and LSTMAE approach. It is essential to observe that the comparisons are for just reference, as several researchers have utilized diverse volumes of data distributions, sampling techniques, and preprocessing methods. Therefore, a simple evaluation for metrics, such as testing and training time, is generally not suitable. Although the proposed ID system attained enhanced performance for the considered evaluation metrics, it cannot be fascinated that the proposed approach fully outclassed other methods. It is possible to attain an extraordinary level of network security with the HIID approach, which is vigorous, fast, simple, and highly applicable to real-time scenarios.

Table 7: Comparison of existing approaches to ISCX-2012 data

Reference	Approach	DR (%)	False alarm rate (%)
Thi-Thu et al. [28]	FS + DT + Variant of RNN	96.33	NA
Kumar et al. [44]	AMGA2 – NB	43.2	7.0
Tan et al. [47]	MCA + EMD	90.12	7.92
Sally et al. [48]	PLL + NGL	95.31	0.80
Heidarian et al. [62]	SVM	89.6	8.6
Keisuke et al. [63]	IDS using Hadoop	86.2	13
Hamed et al. [64]	RFA bigram Approach	89.6	2.6
Mighan et al. [65]	SAE + Classical classifiers	90.3	9.8
Kumar et al. [66]	Ensemble approach	97.0	2.4
Li et al. [67]	RNN – RBM	93.83	1.98
Our approach	HIIDS	97.52	1.2

5 Conclusion and Future Work

In this article, the HIIDS was developed using the Spark MLlib and LSTMAE deep learning approach, which is an efficient cybersecurity method. We trained the HIIDS using an ISCX-2012 dataset. We implemented the HIIDS using several robust classification algorithms, such as LR and XGB, for anomaly detection at Stage1 and the LSTMAE deep learning technique for misuse detection at Stage 2. The proposed HIIDS, based on DL classification, combines the benefits of both Signature-based (SB) and Anomaly-based (AB) approaches, reducing computational complexity and increasing ID accuracy and DR.

Both conventional ML and LSTMAE deep learning models were evaluated using well-known classification metrics, such as F1 score, Precision, Recall, DR, and accuracy of classification.

We believe that our approach can be expanded to other domains in the future; misuses and anomalies can be recognized in several real-time image data, emphasis on exploring deep learning as a features extraction mechanism to learn knowledgeable data illustrations in case of other anomaly detection issues in modern real-time datasets.

Funding Statement: This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-2016-0-00465) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] X. C. Shen, J. X. Du and F. Zhang, "An intrusion detection system using a deep neural network with gated recurrent units," *IEEE Access*, vol. 6, pp. 48697–48707, 2018.

[2] K. Liu, S. Xu, G. Xu, M. Zhang, D. Sun *et al.,* "A review of android malware detection approaches based on machine learning," *IEEE Access*, vol. 8, pp. 124579–124607, 2020.

[3] M. A. Khan and J. Kim, "Toward developing efficient Conv-AE-based intrusion detection system using the heterogeneous dataset," *Electronics*, vol. 9, no. 11, pp. 1–17, 2020.

[4] J. Kim and H. Kim, "An effective intrusion detection classifier using long short-term memory with gradient descent optimization," in *Proc. Platform Technology and Service (Plat Con)*, Busan, South Korea, pp. 1–5, 2017.

[5] G. E. Hinton, S. Osindero and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[6] H. Alqahtani, I. H. Sarker, A. Kalim, S. M. Hossain, S. Ikhlaq *et al.,* "Cyber intrusion detection using machine learning classification techniques," in *Proc. Computing Science, Communication and Security*, Gujarat, India, pp. 121–131, 2020.

[7] N. Kaloudi and L. Jingyue, "The AI-based cyber threat landscape: A survey," *ACM Computing Surveys*, vol. 53, no. 1, pp. 1–34, 2020.

[8] B. Li, Y. Wu, J. Song, R. Lu, T. Li *et al.,* "Deep Fed: Federated deep learning for intrusion detection in industrial cyber-physical systems," *EEE Transactions on Industrial Informatics*, vol. 1, pp. 1–10, 2020.

[9] M. A. Ferrag, L. Maglaras, S. Moschoyiannis and H. Janicke, "Deep learning for cybersecurity intrusion detection approaches datasets and comparative study," *Journal of Information Security and Applications*, vol. 50, pp. 1–19, 2019.

[10] W. Zong, Y. W. Chow and W. Susilo, "Interactive three-dimensional visualization of network intrusion detection data for machine learning," *Future Generation Computer Systems*, vol. 102, no. 4, pp. 292–306, 2020.

[11] C. Khammassi and S. Krichen, "A GA-LR wrapper approach for feature selection in network intrusion detection," *Computers and Security*, vol. 70, no. 2, pp. 255–277, 2017.

[12] N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective*, vol. 25, no. 13, pp. 18–31, 2016.

[13] J. A. Khan and N. Jain, "A survey on intrusion detection systems and classification techniques," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 2, no. 5, pp. 202–208, 2016.

[14] N. Farnaaz and M. A. Jabbar, "Random forest modeling for network intrusion detection system," *Procedia Computer Science*, vol. 89, no. 1, pp. 213–217, 2016.

[15] S. S. S. Sindhu, S. Geetha and A. Kannan, "Decision tree-based lightweight intrusion detection using a wrapper approach," *Expert Systems with Applications*, vol. 39, no. 1, pp. 129–141, 2012.

[16] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cybersecurity intrusion detection," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016.

[17] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P. A. Manzagol *et al.,* "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, no. 12, pp. 3371–3408, 2010.

[18] G. Kim, S. Lee and S. Kim, “A novel hybrid intrusion detection method integrating anomaly detection with misuse detection,” *Expert Systems with Applications*, vol. 41, no. 4, pp. 1690–1700, 2014.

[19] N. Paulauskas and J. Auskalnis, “Analysis of data pre-processing influence on intrusion detection using NSL-KDD dataset,” in *Proc. Open Conf. of Electrical Electronic and Information Sciences IEEE*, Vilnius, Lithuania, pp. 1–5, 2017.

[20] S. Zaman and F. Karray, “Features selection for intrusion detection systems based on support vector machines,” in *Proc. of Consumer Communications and Networking IEEE*, Las Vegas, NV, USA, pp. 1–8, 2009.

[21] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, M. Ghogho *et al.,* “Deep learning approach for network intrusion detection in software-defined networking,” in *Proc. Int. Conf. on Wireless Networks and Mobile Communications*, Fez, Morocco, pp. 258–263, 2016.

[22] M. A. Qatf, Y. Lasheng, M. A. Habib and K. A. Sabahi, “Deep learning approach combining sparse autoencoder with SVM for network intrusion detection,” *IEEE Access*, vol. 6, pp. 52843–52856, 2018.

[23] F. Farahnakian and J. Heikkonen, “A deep auto-encoder based approach for an intrusion detection system,” in *Proc. 20th Int. Conf. on Advanced Communication Technology*, Chuncheon-si Gangwon-Do, South Korea, pp. 178–183, 2018.

[24] S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han *et al.,* “Enhanced network anomaly detection based on deep neural networks,” *IEEE Access*, vol. 6, pp. 48231–48246, 2018.

[25] S. Bandyopadhyay, R. Chowdhury, A. Roy and B. Saha, “A step forward to revolutionize intrusion detection system using deep convolution neural network,” *Preprints*, vol. v1, pp. 1–13, 2020.

[26] M. A. Albahar and M. Binsawad, “Deep autoencoders and feedforward networks based on a new regularization for anomaly detection,” *Security and Communication Networks*, vol. 2020, no. 8, pp. 1–9, 2020.

[27] M. Monshizadeh, V. Khatri, B. G. Atli, R. Kantola and Z. Yan, “Performance evaluation of a combined anomaly detection platform,” *IEEE Access*, vol. 7, pp. 100964–100978, 2019.

[28] L. T. Thu, Y. Kim and H. Kim, “network intrusion detection based on novel feature selection model and various recurrent neural networks,” *Applied Sciences*, vol. 9, no. 7, pp. 1–29, 2019.

[29] S. N. Mighan and M. Kahani, “Deep learning-based latent feature extraction for intrusion detection,” in *Proc. Electrical Engineering (ICEE), Iranian Conf. on IEEE*, Mashhad, Iran, pp. 1511–1516, 2018.

[30] W. Wang, Y. Sheng, J. Wang, X. Zeng, X. Ye *et al.,* “HAST-IDS learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection,” *IEEE Access*, vol. 6, pp. 1792–1806, 2017.

[31] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang *et al.,* “Deep learning and Its applications to machine health monitoring: A survey,” *arXiv preprint arXiv*, vol. 14, no. 8, pp. 1–14, 2016.

[32] S. M. Erfani, S. Rajasegarar, S. Karunasekera and C. Leckie, “High dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning,” *Pattern Recognition*, vol. 58, no. 7, pp. 121–134, 2016.

[33] U. Fiore, F. Palmieri, A. Castiglione and A. D. Santis, “Network anomaly detection with the restricted Boltzmann machine,” *Neurocomputing*, vol. 122, no. 3, pp. 13–23, 2013.

[34] Z. Wang, “The applications of deep learning on traffic identification,” *Black Hat Tech*, vol. 24, no. 11, pp. 1–10, 2013.

[35] Q. Niyaz, W. Sun, A. Y. Javid and M. Alam, “A deep learning approach for network intrusion detection system,” in *Proc. 9th EAI Int. Conf. on Bio-Inspired Information and Communications Technologies*, New York City, USA, pp. 21–26, 2016.

[36] C. Yin, Y. Zhu, J. Fei and X. He, “A deep learning approach for intrusion detection using recurrent neural networks,” *IEEE Access*, vol. 5, pp. 21954–21961, 2017.

[37] M. A. khan, M. Karim and Y. Kim, “A scalable and hybrid intrusion detection system based on the convolutional-LSTM network,” *Symmetry*, vol. 11, no. 4, pp. 1–14, 2019.

[38] K. Alrawashdeh and C. Purdy, “Toward an online anomaly intrusion detection system based on deep learning,” in *Proc. Int. Conf. on Machine Learning and Applications, IEEE*, Anaheim, CA, USA, pp. 195–200, 2016.

[39] N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, 2018.

[40] J. Kim, S. Y. N. Shin and S. H. Kim, "Method of intrusion detection using deep neural network," in *Proc Int. Conf. on Big Data and Smart Computing*, Jeju, South Korea, pp. 313–316, 2017.

[41] G. Wang, J. Hao, J. Ma and L. Huang, "A new approach to intrusion detection using artificial neural networks and fuzzy clustering," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6225–6232, 2010.

[42] S. Mukkamala, A. H. Sung and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," *Journal of Network and Computer Applications*, vol. 28, no. 2, pp. 167–182, 2005.

[43] M. Kakavand, N. Mustapha, A. Mustapha and M. T. Abdullah, "Effective dimensionality reduction of payload-based anomaly detection in TMAD model for HTTP Payload," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 8, pp. 3884–3910, 2016.

[44] G. Kumar and K. Kumar, "Design of an evolutionary approach for intrusion detection," *Scientific World Journal*, vol. 2013, no. 962185, pp. 1–14, 2013.

[45] W. Yassin, N. I. Udzir, Z. Muda and M. N. Sulaiman, "Anomaly-based intrusion detection through k-means clustering and naive Bayes classification," in *Proc. of 4th Int. Conf. on Computing and Informatics*, Sarawak, Malaysia, pp. 298–303, 2013.

[46] M. H. Tahir, A. M. Said, N. H. Osman, N. H. Zakaria, P. N. M. Sabri *et al.,* "Oving K-means clustering using discretization technique in network intrusion detection system," in *Proc. of 3rd Int. Conf. on Computer and Information Sciences*, Kuala Lumpur, Malaysia, pp. 248–252, 2016.

[47] Z. Tan, A. Jamdagni, X. He, P. Nanda, R. P. Liu *et al.,* "Detection of denial-of-service attacks based on computer vision techniques," *IEEE Transactions on Computers*, vol. 64, no. 9, pp. 2519–2533, 2015.

[48] H. Sallay, A. Ammar, M. B. Saad and S. Bourouis, "A real-time adaptive intrusion detection alert classifier for high-speed networks," in *Proc. of IEEE 12th Int. Symp. on Network Computing and Applications NCA*, USA, pp. 73–80, 2013.

[49] Y. Zhou, T. A. Mazzuchi and S. Sarkani, "M-AdaBoost-A based ensemble system for network intrusion detection," *Expert Systems with Applications*, vol. 162, pp. 1–15, 2020.

[50] P. Bedi, N. Gupta and V. Jindal, "I-Siam IDS: An improved Siam-IDS for handling class imbalance in network-based intrusion detection systems," *Applied Intelligence*, vol. 1, pp. 1–19, 2020.

[51] F. Thabtah, F. Kamalov, S. Hammoud and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Information Sciences*, vol. 513, no. 3, pp. 429–441, 2020.

[52] A. Shiravi, H. Shiravi, M. Tavallee and A. A. Ghorbani, "Toward developing a systematic approach to generate benchmark datasets for intrusion detection," *Computers and Security*, vol. 31, no. 3, pp. 357–374, 2012.

[53] X. Meng, J. Bradley, B. Yavuz, E. Sparks, S. Venkataraman *et al.,* "MLlib: Machine learning in apache spark," *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1235–1241, 2016.

[54] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust *et al.,* "Apache spark: A unified engine for big data processing," *Communications of the ACM*, vol. 59, no. 11, pp. 56–65, 2016.

[55] S. Hochreiter and J. S. Huber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[56] F. Gers, N. Schraudolph and J. S. Huber, "Learning precise timing with LSTM recurrent networks," *Journal of Machine Learning Research*, vol. 3, pp. 115–143, 2002.

[57] M. A. khan, M. Karim and Y. Kim, "A two-stage big data analytics framework with real-world applications using spark machine learning and long short-term memory network," *Symmetry*, vol. 10, no. 485, pp. 1–22, 2018.

[58] Z. Wu, X. Wang, Y. G. Jiang, H. Ye and X. Xue, "Modeling spatial-temporal clues in a hybrid deep learning framework for video classification," in *Proc. of the 23rd ACM Int. Conf. on Multimedia*, Brisbane, Australia, pp. 461–470, 2015.

[59] D. Tang, B. Qin and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," in *Proc. of the Conf. on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, pp. 1422–1432, 2015.

[60] Y. Fan, X. Lu, D. Li and Y. Liu, "Video-based emotion recognition using Cnn-rnn and c3d hybrid networks," in *Proc. of the 18th ACM Int. Conf. on Multimodal Interaction*, Tokyo, Japan, pp. 445–450, 2016.

[61] K. Vignesh, G. Yadav and A. Sethi, "Abnormal event detection on BMTT-PETS, 2017 surveillance challenge," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, Honolulu, Hawaii, pp. 2161–2168, 2017.

[62] Z. Heidarian, N. Movahedinia, N. Moghim and P. Mahdinia, "Intrusion detection based on normal traffic specifications," *International Journal of Computer Network and Information Security*, vol. 7, no. 9, pp. 32–38, 2015.

[63] K. Kato and K. Vitaly, "Development of a network intrusion detection system using apache Hadoop and spark," in *Proc. Conf. on Dependable and Secure Computing IEEE*, Taipei, Taiwan, pp. 2539–2553, 2017.

[64] T. Y. Hamed, R. Dara and K. C. Stefan, "Network intrusion detection system based on recursive feature addition and bigram technique," *Computers and Security*, vol. 73, no. 3, pp. 137–155, 2018.

[65] S. N. Mighan and M. Kahani, "A novel scalable intrusion detection system based on deep learning," *International Journal of Information Security*, vol. 2020, no. 3, pp. 1–7, 2020.

[66] G. Kumar, "An improved ensemble approach for effective intrusion detection," *Journal of Supercomputing*, vol. 76, no. 1, pp. 275–291, 2020.

[67] C. Li, J. Wang and X. Ye, "Using a recurrent neural network and restricted Boltzmann machines for malicious traffic detection," *Neuro Quantology*, vol. 16, no. 5, pp. 1–10, 2018.